PROJECT SENSORHUB

MATTHIAS MÜLLER
“The ultimate goal with Big Data is to collect, store, find meaning, and extract value from data. [...] As things stand now, the data ecosystem is highly fragmented. Between those who create data and those who could potentially extract value from it sits a labyrinth fraught with complexity, disparity, and miscommunication.” – Cisco IBSG, 2012
OUR VISION

Environmental data at YOUR finger tips

- A data-driven perspective on our natural and built environment
- Democatized access to environmental data
- Inform and educate citizens and organizations to make better, evidence-based decisions
SENSORHUB / OPENSENSORWEB

> 15 networks
> 115K devices
> 340K sensors
> 4.3 Bn measurements
OPENSENSORWEB – FEATURES

• Rich data: A data integration platform for local environmental data for Germany, Europe and the World
• Convenient search and data exploration functions
• Harmonized data formats
• Extremely fast access to time series data
• Modern user interface and APIs
TIME SERIES DATABASE

- High throughput
  - Fast writes (400k inserts / second) *
  - Ultra fast reads (8M measurements per second) *
- High compression ( < 1 byte / measurement)
- Storage of "true" values (0.35 != 0.35000000000000003d)
- Interval-algebra (we don’t generalize time intervals to instants)
- Exact on-the-fly aggregation and generalization
  - SUM, AVERAGE, including MIN/MAX spans
- Runs on commodity hardware, horizontally scalable

* single node, single threaded
INTERACTIVE CHARTS AND MAPS

• Optimized for large amounts of data (or low bandwidth)
• Ultra fast chart interaction with many large time serieses
• We only request what we render
  • No need to query data at 1-minute resolution when you display daily averages or sums
  • Hyrbid server- and client-side clustering and filtering of sensor locations and phenomenons
• Progressive Webapps for Mobile and Desktop
• Areal interpolation in WebGL
ROADMAP

• September 2017:
  • Revised taxonomy of environmental phenomenons
  • Final UI design and interaction
• October 2017: OpenSensorWeb public release
• ~ February 2018:
  • Data acquisition (a bigger data pool)
  • Additional visualisation and data filtering capabilities
• ~ 2018+:
  • Live monitoring (for network operators)
  • Machine learning
  • Crowd sensing
TOWARDS DATA PRODUCTS

• Integrated and harmonized Environmental Sensor Data

• Derived Analyses (densified data, regression models, trends, forecasts)

• Growing data pool for data mining and machine learning projects
WE ARE LOOKING FOR PARTNERS WITH ...

• Integration challenges
  • Interesting (big) data archives!
  • Sensor infrastructures sending live data
• Technical challenges
  • Partners that challenge our technology stack
  • ... or want to combine it with existing data management systems
• Feedback and wishlists